STOCHASTIC ANALYSIS AND APPLICATIONS, 11(2), 133-153 (1993)

ON THE SIMULATION OF EXPECTATIONS OF RANDOM VARIABLES DEPENDING ON A STOPPING TIME

M. B. Alaya

CERMA - ENPC la Courtine
93167 Noisy le grand cedex France,
email : mba@enpc.fr

ABSTRACT

Birkhoff's pointwise ergodic theorem with the shift operator over $[0,1]^{\mathbb{N}}$ yields a new practical method to compute expectations of functionals of stochastic process. Indeed $\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n$ converges to $\mathbb{E}F$, as N converges toward infinity, almost surely. By numerical simulations we will explain the efficiency of this method especially when compared to the classic Monte-Carlo one. It will furthermore be proven that under suitable assumptions a central limit theorem holds. These assumptions are satisfied in most encountered practical problems. It will precisely be fulfilled when $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$ with a stopping time T having a moment of order p, $p > 2$. Moreover, under this assumptions a "weak" law of iterated logarithm applies. Such that :

$$\forall \varepsilon > 0 \qquad |\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F| = o(N^{\frac{-1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad dx^{\otimes \mathbb{N}} a.s..$$

Numerical simulations were processed.

1 INTRODUCTION.

Concerning numerical simulations in large dimension or for random processes, the use of Birkhoff's pointwise ergodic theorem in the case of the shift operator, turns out to be a an efficient method in many aspects, especially when taking into account its implementation on computers (see Bouleau[3]). Therefore it is interesting to study in details the laws that rule the behavior of this method and the rate of convergence for the functions currently encountered in actual simulations.

Let $[0,1]^{\mathbb{N}}$ be the product space of the interval $[0,1]$, with the Lebesgue's measure $\lambda = dx^{\otimes \mathbb{N}}$ and let θ be the shift operator, defined on $[0,1]^{\mathbb{N}}$ as follow

$$\theta(U_1, U_2, \ldots, U_k, \ldots) = (U_2, U_3, \ldots, U_{k+1}, \ldots).$$

It is easy to check that $\theta(\lambda) = \lambda$, where $\theta(\lambda)$ is the image measure of λ by θ. It is well known that the dynamic system $([0,1]^{\mathbb{N}}, \mathcal{B}([0,1]^{\otimes \mathbb{N}}), \lambda, \theta)$ is ergodic (see Krengel[5]). Therefore, it derives from the Birkhoff pointwise ergodic theorem, that :

$$\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n \longrightarrow \mathbb{E}F \qquad \lambda - a.s., \tag{1}$$

for every λ-integrable function F on $[0,1]^{\mathbb{N}}$.

Let us recall two results for the shift operator over $([0,1]^{\mathbb{N}}, \mathcal{B}([0,1]^{\otimes \mathbb{N}}), \lambda, \theta)$ (see. Krengel [5]), we deduce that the rate of convergence can be arbitrarily slow or arbitrarily close to $0(\frac{1}{N})$.

On a one hand : for every sequence $(\alpha_n)_{n\in\mathbb{N}}, \alpha_n > 0, \alpha_n \to 0$, there exist a continuous function F over $[0,1]^{\mathbb{N}}$ of real values such that

$$\frac{1}{\alpha_N}(|\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F|) \longrightarrow \infty. \qquad \lambda - a.s..$$

On the other hand : for every sequence $(c_n)_{n\in\mathbb{N}}, c_n > 0$, increasing toward infinity, with $c_1 \geq 2$, there exist a measurable set A with $P(A) = \frac{1}{2}$, for which

$$\forall N \qquad |\frac{1}{N}\sum_{n=0}^{N-1} 1_A \circ \theta^n - \frac{1}{2}| \leq \frac{c_N}{N}.$$

One goal of this paper is to explain to the reader, with examples inspired by actual problems, the interest of this method. We will study the asymptotic law of $\frac{1}{\sqrt{N}}(\sum_{n=0}^{N-1} F \circ \theta^n - N\mathbb{E}F)$ and determine the rate of convergence for several classes of functions usually employed in actual simulations.

Through out this paper, we will denote :

$$\sigma^2 = \sigma^2(F) = Var(F) + 2\sum_{k=1}^{\infty} cov(F \circ \theta^k, F),$$

if the series $\sum_{k=1}^{\infty} cov(F \circ \theta^k, F)$ is converging, and $\sigma^2 = +\infty$ otherwise. Furthermore, let set :

$$Var(F) = \mathbb{E}(F - \mathbb{E}F)^2 \quad , \quad cov(F, G) = \mathbb{E}((F - \mathbb{E}F)(G - \mathbb{E}G))$$

for all F,G $\in L^2([0,1]^{\mathbb{N}})$.

The first part of this paper is dedicated to the purely numerical aspects of the method. Indeed the particularities of its implementation on computer will be discussed and its efficiency will be tested by simulation, especially in comparaison with the classic Monte-Carlo method. In the second part the theoretical results, base on hypothesis that are usually satisfied during simulation, will be presented. It will furthermore be proven that if T is a stopping time having a moment of order p, $p > 2$ and if F is a function of L^2, $\mathcal{F}_T$-measurable with null integral, then the central limit theorem applies and a "weak" law of iterated logarithm holds. Indeed, on a one hand :

$$\frac{1}{\sigma\sqrt{N}}\sum_{n=0}^{N-1} F \circ \theta^n \xrightarrow{\mathcal{L}} \mathcal{N}(0;1). \tag{2}$$

With $\mathcal{N}(0;1)$ the standard normal distribution and $\xrightarrow{\mathcal{L}}$ means convergence in distribution. On the other hand :

$$\forall \varepsilon > 0 \qquad \left|\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F\right| = o(N^{\frac{-1}{2}}(\log(N))^{\frac{3}{2}+\epsilon}) \qquad \lambda - p.s.. \tag{3}$$

The third part illustrates with numerical simulations the theoretical results obtained in the second one. Furthemore, we will note that usual law of iterated logarithm is verified by simulation.

2 NUMERICAL ASPECT OF THE METHOD.

The modelling of random phenomena has often to deal with discret time processes such as : particles motion, the discretization of stochastic differential equations, etc. Very often these processes are Markov chains. We will describe how to implement such a chain with the shift method.

Let X_n a Markov chain defined by :

$$X_0 = x, \qquad X_{n+1} = H(X_n, n, h(U_{nd+1}, \cdots, U_{(n+1)d})),$$

where U_n are i.i.d. random variables with the uniform distribution on the interval [0,1]. Therefore the process $(X_n)_{n\in\mathbb{N}}$ admits a representation on $([0,1]^{\mathbb{N}}, dx^{\otimes\mathbb{N}})$.

Let T be the reaching time of the Borel set A, $T = inf\{n \geq 1; X_n \in A\}$. We associate to the stopping time T the class of functions defined by $F = G(X_T, T)$. The expectation $\mathbb{E}G(X_T, T)$ can be evaluated by Birkhoff's theorem, because

$$\mathbb{E}G(X_T, T) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} G(X_T, T) \circ \theta^{nd} \qquad \lambda - a.s., \tag{4}$$

as soon as $G(X_T, T) \in L^1$. The practical implementation of this algorithm calls upon the notion of pointer (or any other process equivalent to stock management).

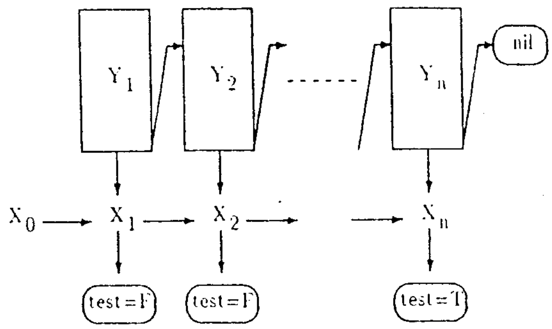

FIG 1.

In order to apply equation (3), $G(X_T, T)$ has to be evaluated at the successive points $U, \theta^d(U), \theta^{2d}(U), \ldots$ of $[0,1]^{\mathbb{N}}$, each of them being a [0,1]-valued sequence

$$
\begin{aligned}
U &= (U_1, U_2, \cdots, U_d, \cdots, U_k, \cdots) \\
\theta^d(U) &= (U_{d+1}, U_{d+2}, \cdots, U_{2d}, \cdots, U_{d+k}, \cdots) \\
&\vdots
\end{aligned}
$$

Therefore, if the stopping time T is almost surely finite, which is what we assume here, the evaluation of $G(X_T, T)$ depends only on a finite number of coordinates along each trajectory $U = (U_1, U_2, \cdots, U_k, \cdots)$. The sequence U already computed until the indice $k = dT(U)$ where $T(U)$ is determined by the test

$$
< test > \left\{ \begin{array}{ll} X_n \text{ still outside } A & \Rightarrow \quad n < T(U) \\ X_n \in A \text{ for the first time} & \Rightarrow \quad n = T(U) \end{array} \right.
$$

Let us assume that $G(X_T, T)$ was evaluated at the point $U = (U_1, U_2, \cdots)$, and that the value of intermediate terms, $Y_n = h(U_{(n-1)d+1}, \cdots, U_{nd})$, $1 \le n \le T(U)$, were placed in pointers as shown in FIG 1. The simulation of $G(X_T, T)$ along the trajectory $\theta^d(U) = (U_{d+1}, U_{d+2}, \cdots)$ will be easier because the intermediate evaluation of the Y_i's associated to $\theta^d(U)$ is partially done. Indeed, either $T(\theta^d(U)) < T(U)$ and no other computation is required, or the chain has just to be lengthen from $T(U)$ to $T(\theta^d(U))$. If the second case occurs the FIG 2 holds .

In this diagrams the arrows represent the evaluation of :

$$X_{n+1} = H(X_n, n, h(U_{nd+1}, \cdots, U_{(n+1)d})) = H(X_n, n, Y_{n+1}).$$

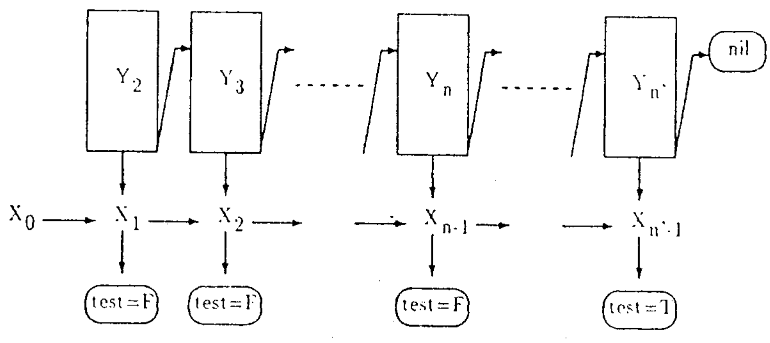

FIG 2.

It is clear that the performances of this method derive emanate from the storage box that permits to avoid redoing the partial products of the variables Y_i. Of course this storage has to be done during the lengthening of the sequence when this latter is needed as explained above.

In the rest of this part we will test the efficiency of this method, we will compare it to the classical method of Monte-Carlo and we will emphasize through an example the saving of C.P.U time, as well as its saving in random number. Indeed one way to simulate the real random walk of the standard normal distribution, issued from the origin, is to set

$$X_0 = 0, \quad X_{n+1} = X_n + \sqrt{-2\ln U_{2n+1}}cos(2\pi U_{2n+2}),$$

with $(U_n)_{n\in N}$ a sequence of i.i.d. random variables with uniform distribution on $[0,1]$. Let T be the stopping defined by $T = \inf\{n \geq 1, X_n \notin [-10,10]\}$.

- For $F = X_T$.

TABLE 1.

Iteration number	C.P.U. time		Random number	
	Shift	Monte-Carlo	Shift	Monte-Carlo
1000	1.3	18.0	2876	220232
5000	5.2	93.1	10214	1132330
10000	10.5	199.5	20610	2260466
50000	59.0	916.4	100664	11209942
100000	118.8	2078.1	200108	22440720

- For $F = T$.

TABLE 2.

Iteration number	C.P.U. time		Random number	
	Shift	Monte-Carlo	Shift	Monte-Carlo
1000	1.3	18.2	2876	220232
5000	5.3	94.3	10214	1132330
10000	10.8	187.2	20610	2260466
50000	60.4	936.4	100664	11209942
100000	123.1	1854.1	200108	22440720

3 ASYMPTOTIC RESULTS ON THE SHIFT.

In general, we will consider a $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping time T and a $\mathcal{F}_T$-measurable function F, with $\mathcal{F}_n = \sigma(U_1, \cdots, U_n)$.

Proposition 1 .

If T has a moment of order $p > 2$ then for every $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$, we have

$$| < F \circ \theta^k, F > | \le \frac{\int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda (\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}},$$

subsequently the series $\sigma^2(F)$ is absolutely converging.

Proof : We have

$$< F \circ \theta^k, F > = < F \circ \theta^k, F \cdot 1_{\{T \le k\}} > + < F \circ \theta^k, F \cdot 1_{\{T > k\}} > .$$

From the fact that $F \circ \theta^k$ is $\sigma(U_{k+1}, \cdots)$-measurable and that $F \cdot 1_{\{T \le k\}}$ is $\sigma(U_1, \cdots, U_k)$-measurable, we deduce that $F \circ \theta^k$ and $F \cdot 1_{\{T \le k\}}$ are independent. Therefore

$$< F \circ \theta^k, F \cdot 1_{\{T \le k\}} > = \int_{[0,1]^{\mathbb{N}}} F \circ \theta^k d\lambda \int_{[0,1]^{\mathbb{N}}} F \cdot 1_{\{T \le k\}} d\lambda = 0,$$

furthermore

$$\begin{aligned} | < F \circ \theta^k, F > | &= | < F \circ \theta^k, F \cdot 1_{\{T > k\}} > | \\ &= | < F \circ \theta^k \cdot 1_{\{T > k\}}, F > | \\ &\le \left(\int_{[0,1]^{\mathbb{N}}} |F \circ \theta^k \cdot 1_{\{T > k\}}|^2 d\lambda\right)^{\frac{1}{2}} \cdot \left(\int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda\right)^{\frac{1}{2}}. \end{aligned} \quad (5)$$

By noticing that $F \circ \theta^k$ is $\sigma(U_{k+1}, \cdots)$-measurable, and that $\{T > k\}$ belongs to $\sigma(U_1, \cdots, U_k)$, it appears that $F \circ \theta^k$ and $1_{\{T>k\}}$ are independent. Furthermore

$$\begin{aligned} \int_{[0,1]^{\mathbb{N}}} |F \circ \theta^k \cdot 1_{\{T>k\}}|^2 d\lambda &= \int_{[0,1]^{\mathbb{N}}} |F \circ \theta^k|^2 d\lambda \cdot \int_{[0,1]^{\mathbb{N}}} 1_{\{T>k\}} d\lambda \\ &= \int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda \cdot \mathbb{P}(T > k), \end{aligned}$$

Therefore

$$| < F \circ \theta^k, F > | \leq \int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda \cdot (\mathbb{P}(T > k))^{\frac{1}{2}}.$$

because $T \in L^p$, it follows then

$$| < F \circ \theta^k, F > | \leq \frac{C}{k^{\frac{p}{2}}}$$

with

$$C = (\int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda)(\mathbb{E}T^p)^{\frac{1}{2}},$$

which concludes the proof.

∎

Remarks:

1/From the proof, we notice that the condition

$$\sum_{k=1}^{\infty} \mathbb{P}(T > k)^{\frac{1}{2}} < \infty, \tag{6}$$

is actually sufficient to get the series $\sigma^2(F)$ absolutely converging.

2/By applying *Hölder* inequality instead of of Cauchy-Schwartz in equation (5) , we obtain a more precise result; indeed if T satisfies the following property

$$\sum_{k=1}^{\infty} \mathbb{P}(T > k)^{\frac{1}{p}} < \infty,$$

then, for every $F \in L^q([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, $q \geq p$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$ and $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p < \infty$, we have

$$| < F \circ \theta^k, F > | \leq (\int_{[0,1]^{\mathbb{N}}} |F|^p d\lambda)^{\frac{1}{p}} \cdot (\int_{[0,1]^{\mathbb{N}}} |F|^q d\lambda)^{\frac{1}{q}} \cdot \mathbb{P}(T > k)^{\frac{1}{p}}.$$

Especially, if $p = 1$, we obtain that if T is integrable and F is a bounded $\mathcal{F}_T$-measurable function then

$$| < F \circ \theta^k, F > | \leq ||F||_\infty (\int_{[0,1]^{\mathbb{N}}} |F| d\lambda) \cdot \mathbb{P}(T > k)^{\frac{1}{p}}.$$

3/Moreover, the result that the series σ^2 is absolutely converging also holds for the wider class of functionals F that can be approximated by a sequence F_k, such that F_k is $\sigma(U_1, \cdots, U_k)$-measurable and $\sum_{k=1}^{\infty} ||F - F_k||_2 < \infty$. In fact, under this assumptions we have (see thesis, reference[1]) :

$$| < F \circ \theta^k . F > | \leq ||F||_2 ||F - F_k||_2.$$

For the reader's convenience we will prove a second proposition which will be used to establish a central limite theorem and to evalute the rate of convergence of this method.

Proposition 2 *Let $F \in L^2$.*

$$\textit{If} \qquad \sigma^2(F) < \infty \qquad \textit{then}$$

$$\lim_{N \to \infty} \frac{1}{N} Var(\sum_{n=0}^{N-1} F \circ \theta^n) = \sigma^2. \tag{7}$$

Proof : Since, one can substitute $F - \mathbb{E}F$ for F in the previous equation, we will assume that F has a null integral. Then, we have :

$$\begin{aligned}
\int_{[0,1]^{\mathbb{N}}} |\sum_{n=0}^{N-1} F \circ \theta^n|^2 d\lambda &= \sum_{k,l \leq N} < F \circ \theta^k, F \circ \theta^l > \\
&= N||F||_2^2 + 2\sum_{l=1}^{N} (\sum_{k=1}^{N-l} < F \circ \theta^k, F >) \\
&= N\sigma^2 - 2\sum_{l=1}^{N} (\sum_{k=N-l+1}^{N} < F \circ \theta^k, F >) \\
&\quad -2\sum_{l=1}^{N} (\sum_{k=N+1}^{\infty} < F \circ \theta^k, F >) \\
&= N\sigma^2 - 2\sum_{k=1}^{N} k < F \circ \theta^k, F > -2N \sum_{k=N+1}^{\infty} < F \circ \theta^k, F >
\end{aligned}$$

Therefore

$$\frac{1}{N} \int_{[0,1]^{\mathbb{N}}} |\sum_{n=0}^{N-1} F \circ \theta^n|^2 d\lambda = \sigma^2 - \frac{2}{N} \sum_{k=1}^{N} k < F \circ \theta^k, F > -2 \sum_{k=N+1}^{\infty} < F \circ \theta^k, F > .$$

It derives from Kronecker lemma (see for example [6]) the second term, to the right of the equality, converges towards zero. The third term, converges to zero, as a remainder of convergent sequence. Therefore

$$\lim_{N\to\infty}\frac{1}{N}\int_{[0,1]^{\mathbb{N}}}|\sum_{n=0}^{N-1}F\circ\theta^n|^2 d\lambda = \sigma^2$$

∎

Remarks:

1/The functions depending on a finite number of variables, called cylindric functions, verify the condition $\sigma^2 < \infty$. Indeed, if F depends only on the first N coordinates, Then :

$$\sigma^2(F) = \|F\|_2^2 + 2\sum_{k=1}^{N-1} < F\circ\theta^k . F > .$$

2/The previous result is valid for every strong mixing transform :
One recalls that a transform is strong mixing if

$$\lim_{k\to\infty} < f\circ\tau^k, f >= (\int_X f d\mu)^2 \qquad \forall f \in L^2(X, \mathcal{A}, \tau, \mu).$$

where $(X, \mathcal{A}, \tau, \mu)$, is a given dynamic system.

3.1 Convergence In Distribution.

In this part we will prove the central limit theorem for functions depending on a stopping time. Thus we will have the following result :

Theorem 1 .
If T have a moment of order $p > 2$ then for every $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$ and $\sigma^2(F) > 0$ we have

$$\frac{1}{\sigma\sqrt{N}}\sum_{n=0}^{N-1} F\circ\theta^n \xrightarrow{\mathcal{L}} \mathcal{N}(0;1), \tag{8}$$

where $\mathcal{N}(0;1)$ is the standard normal distribution.

In order to establish the central limit theorem we will now compute the limits of $\sigma^2(F.1_{T\leq l})$ and $\sigma^2(F.1_{T>l})$ when l goes to ∞. Indeed if denote for every $l \in \mathbb{N}$ that :

$$\sigma_l^2 := \sigma^2(F.1_{T\leq l}) = Var(F.1_{T\leq l}) + 2\sum_{k=1}^{\infty} cov(F.1_{T\leq l}\circ\theta^k, F.1_{T\leq l})$$

and

$$\tau_l^2 := \sigma^2(F.1_{T>l}) = Var(F.1_{T>l}) + 2\sum_{k=1}^{\infty} cov(F.1_{T>l}\circ\theta^k, F.1_{T>l}),$$

we have the following results :

Lemma 1 :
If T has a moment of order $p > 2$ then for every $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$, we have

$$\lim_{l \to \infty} \sigma_l^2 = \sigma^2. \tag{9}$$

Proof:
Indeed

$$\sigma_l^2 = \sigma^2(F.1_{T \leq l}) = Var(F.1_{T \leq l}) + 2\sum_{k=1}^{\infty} cov(F.1_{T \leq l} \circ \theta^k, F.1_{T \leq l}).$$

It derives from proposition 1 that :

$$\begin{aligned}
|cov(F.1_{T \leq l} \circ \theta^k, F.1_{T \leq l})| &\leq \int_{[0,1]^{\mathbb{N}}} |F.1_{T \leq l} - \mathbb{E}F.1_{T \leq l}|^2 d\lambda \frac{(\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}} \\
&\leq 2\int_{[0,1]^{\mathbb{N}}} |F.1_{T \leq l}|^2 d\lambda \frac{(\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}} \\
&\leq 2\int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda \frac{(\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}}.
\end{aligned}$$

Hence σ_l^2 is a absolutely convergent sequence uniformly with respect to l. Since each term of the serie converges towards $< F \circ \theta^k, F >$, $\lim_{l \to \infty} \sigma_l^2 = \sigma^2$.

∎

Lemma 2 :
If T has a moment of order $p > 2$ then for every $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$, we have

$$\lim_{l \to \infty} \tau_l^2 = 0. \tag{10}$$

Proof:
Indeed

$$\tau_l^2 = \sigma^2(F.1_{T > l} - \mathbb{E}F.1_{T > l}) = Var(F.1_{T > l}) + 2\sum_{k=1}^{\infty} cov(F.1_{T > l} \circ \theta^k, F.1_{T > l}).$$

The same proof as in the above lemma 1 yields :

$$|cov(F.1_{T > l} \circ \theta^k, F.1_{T > l})| \leq 2\int_{[0,1]^{\mathbb{N}}} |F|^2 d\lambda \frac{(\mathbb{E}T^p)^{\frac{1}{2}}}{k^{\frac{p}{2}}} \tag{11}$$

Therefore τ_l^2 is an absolutely convergent sequence uniformly over l. Since each term converges towards 0, $\lim_{l \to \infty} \tau_l^2 = 0$.

∎

Let us prove now the theorem.

Proof:

Let $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, for every $l \in \mathbb{N}$ we have

$$F = (F.1_{T\leq l} - \mathbb{E}F.1_{T\leq l}) + (F.1_{T>l} - \mathbb{E}F.1_{T>l}).$$

Then

$$\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} F\circ\theta^n = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}(F.1_{T\leq l} - \mathbb{E}F.1_{T\leq l})\circ\theta^n + \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}(F.1_{T>l} - \mathbb{E}F.1_{T>l})\circ\theta^n.$$

Since the function $F.1_{T\leq l} - \mathbb{E}F.1_{T\leq l}$ is $\sigma(U_1,\cdots,U_l)$ measurable, the first term on the right of the equality converges in distribution towards $\mathcal{N}(0,\sigma_l^2)$, for every $l \in \mathbb{N}$, with

$$\sigma_l^2 = \sigma^2(F.1_{T\leq l}) = \|F.1_{T\leq l}\|_2^2 + 2\sum_{k=1}^{\infty} cov(F.1_{T\leq l}\circ\theta^k, F.1_{T\leq l}).$$

From lemma 1, one has the convergence in distribution of $\mathcal{N}(0,\sigma_l^2)$ towards $\mathcal{N}(0,\sigma^2)$. Therefore it suffices to show that :

$$\lim_{l\to\infty}\limsup_{N\to\infty}\mathbb{P}\{|\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}(F.1_{T>l} - \mathbb{E}F.1_{T>l})\circ\theta^n| \geq \varepsilon\} = 0$$

for every positive ε. From the Bienaymé-Tchebychev's inequality

$$\mathbb{P}\{|\frac{1}{\sqrt{N}}\sum_{n=0}^{N}(F.1_{T>l} - \mathbb{E}F.1_{T>l})\circ\theta^n| \geq \varepsilon\} \leq \frac{\int_{[0,1]^{\mathbb{N}}} |\sum_{n=0}^{N}(F.1_{T>l} - \mathbb{E}F.1_{T>l})\circ\theta^n|^2}{N\varepsilon^2}.$$

Using the second proposition and letting N tend towards infinity, yields :

$$\limsup_{N\to\infty}\mathbb{P}\{|\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}(F.1_{T>l} - \mathbb{E}F.1_{T>l})\circ\theta^n| \geq \varepsilon\} \leq \frac{\tau_l^2}{\varepsilon^2}.$$

The proof is achieved simply by applying lemma 2. ∎

Remark :

1/Due to the second remark bellow the proposition 1, one can prove the central limit theorem, for a more large class of functions, by the use of same argument. Actually, if T satisfies the property $\sum_{k=1}^{\infty}\mathbb{P}(T > k)^{\frac{1}{p}} < \infty$, then, for every $F \in L^k([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, $k = \sup(q,2)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$ and $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p < \infty$, we have

$$\frac{1}{\sigma\sqrt{N}}\sum_{n=0}^{N-1} F\circ\theta^n \xrightarrow{\mathcal{L}} \mathcal{N}(0;1).$$

2/Moreover, the central limit theorem also holds for a wider class of functionals F that can be approximated by a sequence F_k, such as F_k is $\sigma(U_1, \cdots, U_k)$-measurable and $\sum_{k=1}^{\infty} \|F - F_k\|_2 < \infty$.(See reference[1]) for a proof.

3.2 Rate Of Convergence.

Let us start first with some definitions.
Let (X, μ) be a measured space. We consider a family of functions $(S(M, N, \cdot))_{M,N \in \mathbb{N}}$, that belongs to $L^p(X, \mu)$ and such that $S(M, N, x) \geq 0$ for every M and N in $\mathbb{N}$ and $x \in X$. We finally suppose that S(M,0,·)=0 for every M in $\mathbb{N}$ and that

$$S(M, N, \cdot) \leq S(M, N', \cdot) + S(M + N', N - N', \cdot) \tag{12}$$

for every $M, N, N' \in \mathbb{N}$ and $0 \leq N' \leq N$.
Given this definitions, the following resul is proven.

Theorem 2

$$\textit{If} \qquad \int_X S(M, N, x)^p d\mu = O(\Psi(N)) \qquad \textit{uniformly over } M, \tag{13}$$

where $\frac{\Psi(N)}{N}$ *is a non decreasing function, that for every* $\varepsilon > 0$ *we have :*

$$S(0, N, x) = o(\Psi(N)(\log(N))^{p+1+\varepsilon})^{\frac{1}{p}} \qquad \mu - a.s. \text{ in } X.$$

Proof : This result has been shown in Gál and Koksma's paper "Sur l'ordre de grandeur des fonctions sommables", under the hypothesis that X is a measurable part of an euclidien space. However this hypothesis is not used in the proof.
See reference [4], (theorem 3, page 646) as well as (definition, page 640).

∎

Remark:
We usually use this theorem in the case Ψ(N)=N and $p = 2$, this gives the following result :

$$\text{If} \qquad \int_X S(M, N, x)^2 dx = O(N) \qquad \text{uniformly over M,}$$

then for every $\varepsilon > 0$

$$S(0, N, x) = o(N^{\frac{1}{2}}(\log(N))^{\frac{3}{2}+\epsilon}) \qquad \mu - a.s..$$

Let us now reconsider the dynamic system $([0, 1]^{\mathbb{N}}, \mathcal{B}([0, 1]^{\otimes \mathbb{N}}), \lambda, \theta)$. One can deduce from the previous theorem a strong ergodic result that suggests an estimate of the rate of convergence, within the pointwise ergodic theorem of Birkhoff.

Proposition 3 *Let* $([0,1]^{\mathbb{N}}, \mathcal{B}([0,1]^{\otimes \mathbb{N}}), \lambda, \theta)$ *be the dynamic system associated to the shift operator. Let* F *be a function over* $[0,1]^{\mathbb{N}}$*, such that* $\sigma^2(F) < +\infty$*, then :*

$$\forall \varepsilon > 0 \qquad \frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s.. \tag{14}$$

Proof : This proposition will be deduced from the previous remark. Indeed, let us denote by

$$S(M,N,U) = |\sum_{M}^{M+N-1} F \circ \theta^n(U)|$$

where U is an element of $[0,1]^{\mathbb{N}}$ $\qquad U = (U_1, U_2, \ldots, U_k, \ldots)$.
We will only verify properties (12) and (13). The other ones can be directly deduced from the definition of $S(M,N,.)$.
On one hand : for every N, N' such that $0 \le N' \le N$

$$\begin{aligned} S(M,N,U) &= |\sum_{M}^{M+N-1} F \circ \theta^n(U)| \\ &\le |\sum_{M}^{M+N'-1} F \circ \theta^n(U)| + |\sum_{M+N'}^{M+N-1} F \circ \theta^n(U)| \\ S(M,N,U) &\le S(M,N',U) + S(M+N', N-N', U). \end{aligned}$$

On the other hand : using the result of the previous proposition and the fact that θ preserves the measure λ we have,

$$\int_{[0,1]^{\mathbb{N}}} S(M,N,U)^2 d\lambda = \int_{[0,1]^{\mathbb{N}}} |\sum_{n=0}^{N-1} F \circ \theta^n|^2 \circ \theta^M d\lambda = \int_{[0,1]^{\mathbb{N}}} |\sum_{n=0}^{N-1} F \circ \theta^n|^2 d\lambda = O(N).$$

Next, the hypothesis of the theorem are verified. Therefore for every $\varepsilon > 0$

$$\sum_{n=0}^{N-1} F \circ \theta^n = o(N^{\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s.,$$

furthermore

$$\frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s..$$

∎

Remark :
1/Actually, the previous result remains valid for any strong mixing transform.
Therefore, we have for every function f in $L^2(X,\mu)$, where $(X, \mathcal{A}, \tau, \mu)$ is a given dynamic system with strong mixing transform, such that $Var(f) + 2\sum_{k=1}^{\infty} cov(f \circ \tau^k, f) < \infty$, the following result holds :

$$\forall \varepsilon > 0 \qquad \frac{1}{N}\sum_{k=0}^{N} f \circ \tau^k - \int_X d\mu = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \mu - a.s..$$

Corollary 1 *If T have a moment of order $p > 2$ then for every $F \in L^2([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$ we have*

$$\forall \varepsilon > 0 \qquad \frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s..$$

Proof : This result derived from the previous proposition and the first one. ∎

Due to the remarks of page 5 and by the same argument we deduce the following corollaries.

Corollary 2 *If T satisfies the property $\sum_{k=1}^{\infty} \mathbb{P}(T > k)^{\frac{1}{p}} < \infty$ then, for every $F \in L^k([0,1]^{\mathbb{N}}, \mathcal{F}_T)$, $k = \sup(q, 2)$, such that $\int_{[0,1]^{\mathbb{N}}} F = 0$ and $\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p < \infty$, we have*

$$\forall \varepsilon > 0 \qquad \frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s..$$

∎

Corollary 3 *Let $(F_k)_{k\in\mathbb{N}}$ be a sequence of $L^2([0,1]^{\mathbb{N}}, \mathcal{B}([0,1]^{\otimes\mathbb{N}}), \lambda, \theta)$. One assumes that F_k is $\sigma(U_1, \cdots, U_k)$-measurable and that F_k converges toward F in $L^2([0,1]^{\mathbb{N}}, \mathcal{B}([0,1]^{\otimes\mathbb{N}}), \lambda, \theta)$. Then :*

$$\text{if} \qquad \sum_{k=1}^{\infty} ||F - F_k||_2 < \infty \qquad \text{we have}$$

$$\forall \varepsilon > 0 \qquad \frac{1}{N}\sum_{n=0}^{N-1} F \circ \theta^n - \mathbb{E}F = o(N^{-\frac{1}{2}}(\log(N))^{\frac{3}{2}+\varepsilon}) \qquad \lambda - a.s..$$

∎

Remarks:

1/One knows that if $(X_n)_{n\in\mathbb{N}}$ is a sequence of real random variables that are independent and identically distributed such that $\mathbb{E}X_1^2 < \infty$ and with null integral, then

$$\frac{1}{N}(X_1 + \cdots + X_N) = O(N^{-\frac{1}{2}}(\log\log(N))^{\frac{1}{2}}),$$

the independency hypothesis was weakened in several works (see Berger [2]).

2/It is evident that the estimations of the previous remark, known as the iterated logarithm property, are stronger than the result of proposition 1.

The usefulness of our result is to give an estimate close to the iterated logarithm but under weaker and more natural hypothesis suitable for simulation.

4 NUMERICAL ILLUSTRATION.

To check the validity of the results obtained in the previous section with computer simulations, we have considered the random walks defined as follow :

$$X_0 = x, \qquad X_{n+1} = X_n + h(U_{n+1}^1, \cdots, U_{n+1}^d),$$

where (U_n^i) are i.i.d. random variables with uniform distribution on [0.1]. For such a random walk let us consider the stopping time $T = inf\{n \geq 1; X_n \notin [a,b]\}$ and functions of type $F = G(X_T, T)$. It is clear that T has a moment of order $p > 2$. An estimate of σ is obtained from proposition 2, knowing that :

$$\lim_{N\to\infty} \frac{1}{N} Var(\sum_{n=0}^{N-1} F \circ \theta^n) = \sigma^2.$$

Unfortunately the proof of the previous equation essentially bases on the Kronecker's lemma. So it does not allow the evaluation of the theoretical error. Thus there is no other method available than to wait for a numerical stabilization of this sequence and to take the stable value as the limit : indeed, if we denote Nmax the value from which the first digits of the terms of the sequence are not modified, any more we will take as an estimate of σ the corresponding quantity $\hat{\sigma} = \frac{1}{Nmax} Var(\sum_{n=0}^{Nmax-1} F \circ \theta^n)$.

For every fixed N we simulated $Var(\sum_{n=0}^{N-1} F \circ \theta^n)$ by Monte Carlo's method over 5000 independent trajectories.

4.1 Convergence In Distribution.

In order to numerically verify the convergence in distribution of the sequence

$$A_n(F) = \frac{F + F\circ\theta + \cdots + F\circ\theta^{n-1} - n\mathbb{E}F}{\hat{\sigma}\sqrt{n}},$$

towards the normal distribution $\mathcal{N}(0;1)$, χ^2-distance between $A_n(F)$ et $\mathcal{N}(0;1)$ was computed : to this end, the real line $\mathbb{R}$ was parted into $m = 57$ classes obtained by dicretizing the interval $]-2.8, 2.8[$, with a step 0.1 and considering the residual classe $]-\infty, -2.8[\cup]2.8, +\infty[$.

Then the empirical frequencies $(\frac{n_i}{N})_{1\leq i\leq m}$ were computed over $N = 5000$ independent simulations of $A_n(F)$. Recall that the "Khi-deux-distance" between $A_n(F)$ and $\mathcal{N}(0;1)$ (related to the m above classes) displays as :

$$D_n^N = D_N(A_n(F), \mathcal{N}(0;1)) = \sum_{i=1}^{m} \frac{(n_i - Np_i)^2}{Np_i},$$

with $p_i = \mathbb{P}\{\mathcal{N}(0;1) \in \text{classe } i\}$.

For the large values of N, it is well known that D_n^N may be assumed to have $\chi^2(m-1)$ distribution. Intuitively, if $A_n(F) \xrightarrow{\mathcal{L}} \mathcal{N}(0;1)$, D_n^N must become small as n increases. One way to know if, for a given n, the simulated values of $A_n(F)$ are close enough to a $\mathcal{N}(0;1)$ distribution, is to implement an adequation χ^2-test. Such a decision test displays :

$$\begin{cases} (H_0) & \text{the } A_n(F) \text{ has a } \mathcal{N}(0;1) \text{ distribution} \\ (H_1) & \text{the } A_n(F) \text{ has not a } \mathcal{N}(0;1) \text{ distribution.} \end{cases}$$

The deciding rule is defined at the level α by the critical area
$W = \{(n_1, \cdots, n_m) \text{ tel que } D_n > C_\alpha\}$, where $\mathbb{P}_{H_0}(W) = \alpha$.
If $\alpha = 0.05$ and $m = 57$, we have $C_\alpha = 74.12$. Subsequently the distance D_n may be considered as small as soon as it is lower then C_α.

Examples : Consider the random walk :

$$X_0 = 0, X_{n+1} = X_n + \frac{2U_{n+1} - 1}{2} \text{ and } T = inf\{n \geq 1, X_n \notin [-5, 1]\}.$$

- If $F = X_T$:
  An estimate of $\hat{\sigma}$, $\hat{\sigma} = 23.724$ was computed, as explained above, for $Nmax = 10000$.
  The table below displays the variation of $D_n = D_n^N$ as a function of n.

TABLE 3.

n	100	1000	5000	10000	100000
D_n	31667.18	759.11	113.92	70.39	67.66

The graphic below illustrate the convergence in distribution of $A_n(X_T)$ towards normal distribution $\mathcal{N}(0;1)$. It is actually the histogram of $A_{100000}(X_T)$.

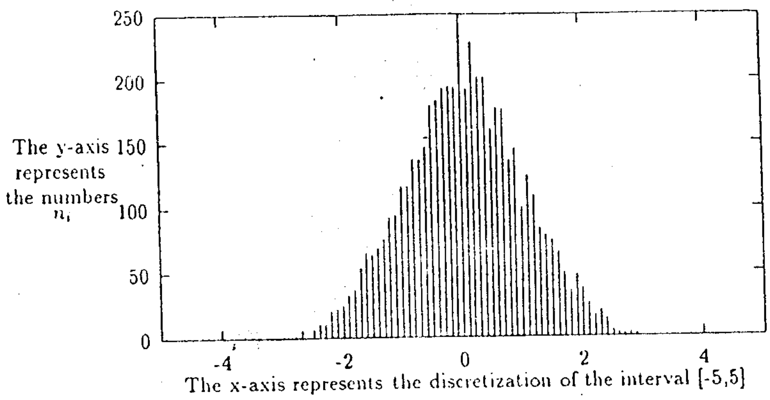

FIG 3.

- Pour $F = T$:

  An estimate of $\hat{\sigma}$, $\hat{\sigma} = 574.281$, was computed, as explained above, for $Nmax = 10000$.

  The table below displays the variation of $D_n = D_n^N$ as a function of n.

TABLE 4.

n	100	1000	5000	10000	100000
D_n	2187.10	258.07	92.10	64.81	72.89

The graphic below illustrate the convergence in distribution of $A_n(X_T)$ towards normal distribution $\mathcal{N}(0;1)$. It is actually the histogram of $A_{100000}(X_T)$.

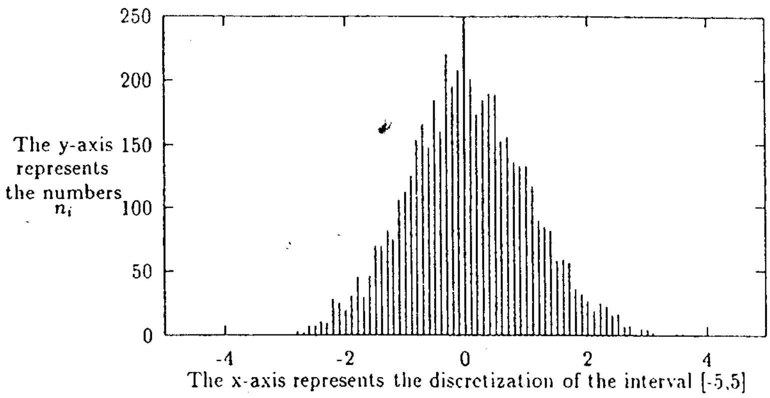

FIG 4.

4.2 Rate Of Convergence.

We wanted to check the validity of the results that we obtained in the preceding paragraph on the rate of convergence and eventually test for the existence of a "classical"law of the iterated logarithmic by simulations.

One simulation of the sample paths of the process

$$B_n(F) = \frac{F + F\circ\theta + \cdots + F\circ\theta^{n-1} - n\mathbb{E}F}{\hat{\sigma}\sqrt{n(\log n)^3}},$$

must therefore confirm the convergence toward zero.

Anyway, a mistake over the evaluation of σ dont prevents us to verify the rate of convergence, since it modify the rapport with a constant factor.

Examples:

Let us consider the random walk

$$X_0 = 0, \qquad X_{n+1} = X_n + \frac{2U_{n+1} - 1}{2},$$

and

$$T = inf\{n \geq 1, X_n \notin [-5, 1]\}.$$

- For $F = X_T$:

  Taking Nmax=10000, we get $\hat{\sigma} = 23.724$

  the following table represent the convergence towards zero of the sequence $B_n(X_T)$:

TABLE 5.

n	1000	10000	100000	200000
$B_n(T)$	0.0249	-0.0032	-0.0358	-0.0231
n	400000	600000	800000	1000000
$B_n(T)$	-0.0238	-0.0014	0.0050	-0.0031

The graphical representation below illustrates the same simulation.

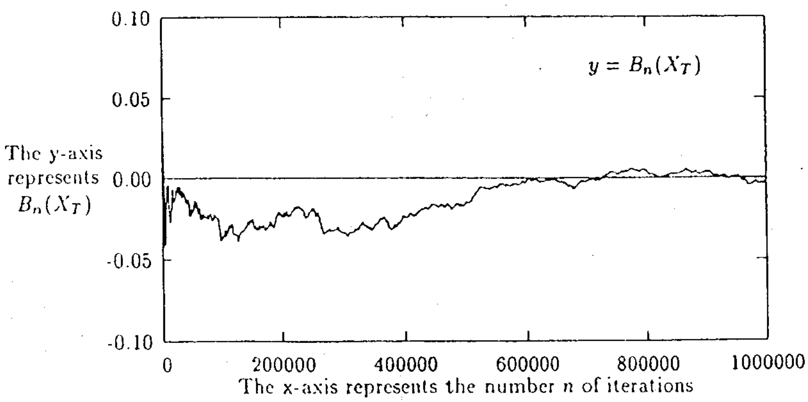

FIG 5.

- For $F = T$:

  by taking Nmax=10000, we obtain $\hat{\sigma} = 574.281$

the following table represent the convergence toward zero of the sequence $B_n(T)$:

TABLE 6.

n	1000	10000	100000	200000
$B_n(T)$	-0.0242	0.0180	0.0383	0.0236
n	400000	600000	800000	1000000
$B_n(T)$	0.0180	0.0067	-0.0037	-0.0057

the graphical representation below illustrates the same simulation.

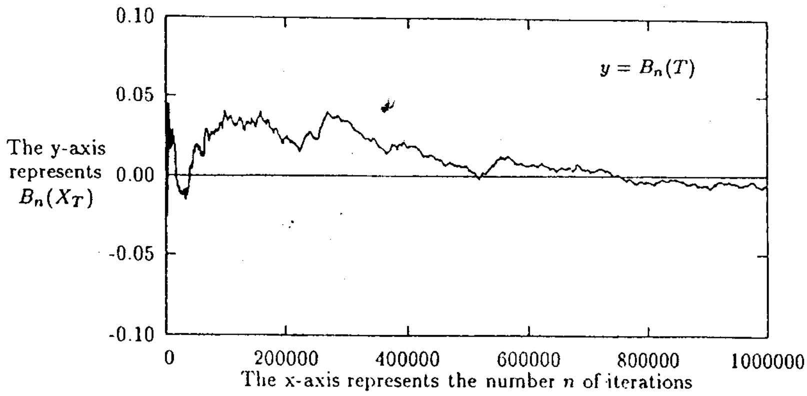

FIG 6.

It seems natural to try to verify experimentally if the law of iterated logarithm is satisfied, this latter was the subject of several previous works (see Berger [2]) and was shown under hypothesis apparently not very well adapted to practical situations which we usually encounter in simulation. In order to do that, we simulated the trajectories of the process :

$$C_n(F) = \frac{F + F \circ \theta + \cdots + F \circ \theta^{n-1} - n\mathbb{E}F}{\hat{\sigma}\sqrt{n \log \log n}}.$$

We clearly notice in the simulations that the plots changed : most of them oscillate between -1 and 1, without converging toward a specific value (see FIG 7 and FIG 8). All the plots of the sequences have been stopped at million time of iteration which seems to be large enough. The fact that all the sample paths remain between -1 and 1, without noticing a convergence toward a specific value is the best confirmation of the law of iterated logarithm that can be numerically given for million steps for this function. This positive result and simulations for other functions suggest that the law of iterated logarithm be satisfied by a sub-class of functions satisfying $\sigma^2 < \infty$. For the examples we studied previously, the following results are obtained :

- For $F = X_T$:
  All the paths of the process that we got have the following shape :

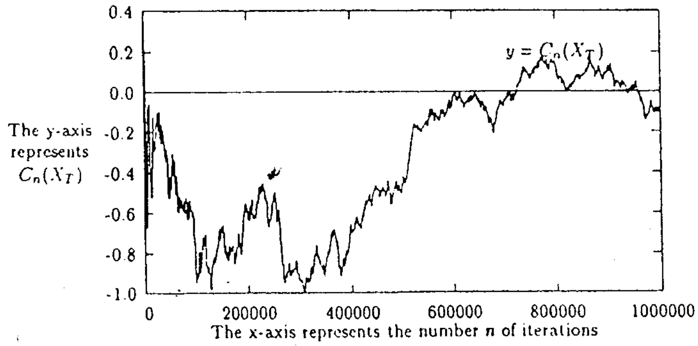

FIG 7

- For $F = T$:
  All the paths of the process that we got have the following shape :

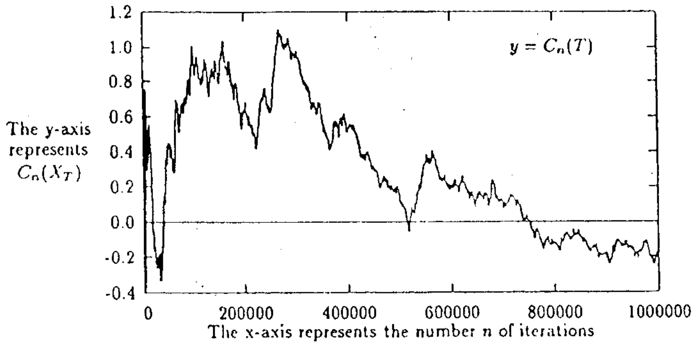

FIG 8.

5 CONCLUSION.

The central limit theorem that we proved for functionals measurable with respect to stopping time and the fact that the rate of convergence is in practice of same order than the law of large number provide a theoritical bases to the practical implementation of the shift method. Let us recall as conclusion that the qualities of this method are essentially : drastic time saving (simulating times are currently cut down by 90% or more in many examples) and even more drastic random generator saving.

REFERENCES

[1] M. B. Alaya : Thèse de l'école Nationale des Ponts et Chaussées, 1992.

[2] E. Berger : An almost sure invariance principle for stationary sequences of Banach space valued random variables, *Probab. th. rel. Fields*, 84, 161-201, 1990.

[3] N. Bouleau : On effective computation of expectations in large or infinite dimension, *Journal of Computational and Applied Mathematics*, 31, 1990, 23-34, North-Holland.

[4] I. S. Gál J. F. Koksma : Sur l'ordre de grandeur des fonctions sommables, *Indigationes Math.*, 12, 1950, 192-207.

[5] U. Krengel : *Ergodic Theorems*, de Gruyter Studies in Mathematics, Berlin, New-York 1985.

[6] J. Neveu : *Martingales à temps discret*, Masson, Paris 1972.